2-9-2015

Screening for Autism Spectrum Disorders in African American Toddlers: Utilizing the Brief Infant Toddler Social Emotional Assessment and Maternal Factors to Improve Classification Accuracy

Rachel Louise Peterman

SCREENING FOR AUTISM SPECTRUM DISORDERS IN AFRICAN AMERICAN TODDLERS: UTILIZING THE BRIEF INFANT TODDLER SOCIAL EMOTIONAL ASSESSMENT AND MATERNAL FACTORS TO IMPROVE CLASSIFICATION ACCURACY

by

Rachel L. Peterman

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Psychology

The University of Memphis

May 2015

Abstract

Peterman, Rachel L. Ph.D. The University of Memphis. May 2015. Screening for Autism Spectrum Disorders in African American Toddlers: Utilizing the BITSEA and Maternal Factors to Improve Classification Accuracy. Major Professor: Beth Meisinger.

Improved diagnostic accuracy and intervention services for toddlers with autism spectrum disorders (ASD) necessitates early identification of children at risk for the disorder. African American and low-income children are at increased risk for delayed diagnosis; however, universal screening may identify those at risk. This study examined the utility of the Brief Infant Toddler Social Emotional Assessment (BITSEA) as an ASD screener in an exclusively African American sample. Moreover, contribution of maternal factors to predict ASD was explored. Participants were selected from the CANDLE (Conditions Affecting Neurocognitive Development and Learning in Early Childhood) study, which is an ongoing study in an urban setting evaluating factors affecting early childhood development. Data were collected from toddlers around 12 and 24 months (i.e., year 1 and year 2) and their mothers. Using items identified by the BITSEA authors, ASD, Dysregulation, and ASD + Dysregulation scales were created for each year and compared to standardized Problem and Competence scales. Internal consistency of BITSEA scales ranged from poor to good. Examination of temporal stability of the scales suggested weak yet significant correlations from both years for all scales. Correlational and ROC curve analyses indicated that the ASD scales outperformed the Problem and Competence scales as indicators of ASD. Cutpoints for the ASD scales produced good sensitivity and specificity at 24 months; however, classification accuracy statistics were lower at 12 months. Regression analyses were employed to examine the contribution of maternal variables and BITSEA ASD scales for predicting ASD risk. Results indicated that the BITSEA ASD score, maternal education, health insurance status, psychological distress, and parenting stress were significantly associated with ASD risk. Toddlers with private insurance, higher BITSEA ASD scores, and higher levels of psychological distress were at greatest risk for ASD compared to others in the study. Results were comparable at both years; however, maternal variables were more predictive at year 1. Overall, these findings suggest that the BITSEA ASD scale can be used to identify African American toddlers at risk for ASD. Additionally, awareness of maternal stress characteristics and ASD symptomology may help identify at risk toddlers who need close monitoring or further evaluation.

Table of Contents

List of Tables

Screening for Autism Spectrum Disorders in African American Toddlers: Utilizing the BITSEA and Maternal Factors to Improve Classification Accuracy

Autism Spectrum Disorders (ASD) are characterized by persistent deficits in social communication and social interaction, along with restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). According to the most recent report by the Centers for Disease Control (CDC), approximately 1 in 68 children are diagnosed with an ASD (2014). Toddlerhood is characterized by the development of social and language skills (American Academy of Pediatrics, 2009) and therefore poses an early opportunity for assessment and intervention for ASD. Indeed, research suggests that ASD can be reliably diagnosed as early as 18 to 24 months (Kim & Lord, 2012; Lord et al., 2006) and interventions for toddlers with ASD have also shown promise (Zwaigenbaum et al., 2009).

As effective interventions for toddlers become more widely available, it is imperative to identify children with ASD at the earliest ages possible. Universal screening can help detect toddlers at risk for ASD and facilitate access to appropriate services. Hence, the American Academy of Pediatrics recommends targeted screenings at both 18 and 24 month well-visits in order to identify toddlers at risk for ASD (Johnson & Meyers, 2007). Although early identification is possible, most children are not diagnosed until about 3 to 4 years of age (Baio, 2012; Howlin & Moore, 1997; Mandell, Novak, & Zubritsky, 2005). Furthermore, some minority groups are at increased risk for delayed diagnosis (Liptak et al., 2008; Mandell, 2002; Mandell et al., 2009; Valicenti-McDermott, Hottinger, Seijo, & Shulman, 2012). African American children under 6 years of age have the lowest rate of diagnosis compared to other ethnicities (Liptak et al., 2008) and they are more likely to receive a misdiagnosis of a behavior disorder instead of ASD

(Mandell, 2007). Additionally, prevalence rates suggest that ASD impacts children similarly across income levels; however, poor children in preschool have the lowest rate of diagnosis (Liptak et al., 2008). Therefore, additional research is needed to examine factors related to the early identification of ASD in African Americans and lower income children to reduce the disparity among underserved populations.

A limited number of reliable and valid instruments are available to assess ASD risk for toddlers (Matson, Wilkins, & Gonzalez, 2008). The Brief Infant Toddler Social Emotional Assessment ([BITSEA] Briggs-Gowan & Carter, 2006) has demonstrated strong psychometric properties as a social-emotional screener to identify behavior problems in community samples (Briggs-Gowan, Carter, Bosson-Heenan, Guyer, & Horwitz, 2006; Briggs-Gowan, Irwin, Wachtel, & Cicchetti, 2004; Briggs-Gowan, Carter, Skuban, & Horwitz, 2001). Recently, two studies have examined the utility of the BITSEA as an ASD screener (Gardner et al., 2013; Karabekiroglua, Briggs-Gowan, Carter, Rodopmu-Armond, & Akbasa, 2010). Gardner et al. (2013) created an ASD scale out of the 17 items identified by the BITSEA authors as indicative of ASD and found that the ASD scale outperformed the Problem and Competence scores in predicting risk for ASD.

The current study seeks to build upon the limited extant research by examining the psychometric properties of the BITSEA in a subsample of African American participants drawn from the same longitudinal study (CANDLE; Conditions Affecting Neurocognitive Development and Learning in Early Childhood) associated with the study by Gardner et al. (2013). A particular need exists to examine the adequacy of ASD screeners with this population due to the risk of delayed diagnosis in African American children. Furthermore, this work seeks to determine whether the predictive validity of

ASD screeners may be enhanced by including additional factors associated with ASD, such as maternal characteristics. For example, higher maternal age and education have been associated with increased risk for ASD (Bilder, Pinborough-Zimmerman, Miller, & McMahon, 2009). In addition, considerable research has found that mothers of children with ASD are more likely to have significantly higher levels of psychological distress (Davis & Carter, 2008; Olsson & Hwang, 2001; Kuusikko-Gauffin et al., 2013) and parenting stress (Davis & Carter, 2008; Estes et al., 2009; Hastings, 2003) when compared to mothers of typical children and children with other developmental disorders. Therefore, the second goal of this study is to examine the contribution of maternal variables (e.g., age, education, psychological distress and parenting stress) with the BITSEA in predicting ASD risk status.

The following literature review describes the characteristics of ASD with emphasis on toddler development. Then, early screening practices for identification of ASD in toddlers are discussed, including previous research on the utility of the BITSEA. Finally, maternal factors associated with ASD are explored as potential contributors to ASD diagnosis.

Literature Review

Characteristics of Autism Spectrum Disorders in Toddlers

The diagnostic criteria for ASD were changed recently with the release of the fifth edition of the Diagnostic and Statistical Manual (APA, 2013) and it subsumes all previous diagnoses of autistic disorder, Asperger's syndrome, and pervasive developmental disorder- not otherwise specified. Based on the new criteria, ASD are characterized by persistent deficits in social communication and social interaction, along with restricted, repetitive patterns of behavior, interests, or activities (APA, 2013).

Additionally, symptoms of ASD must be present in toddlerhood, before 3 years of age. Toddlers rapidly develop a range of social-emotional skills at varying rates. Therefore, the characteristics of ASD are more difficult to identify during this time (Zwaigenbaum et al., 2009). Even so, parent concerns emerge as early as 12 to 18 months (Chawarska et al., 2007; Giacomo & Fombonne, 1998; Howlin & Moore, 1997). Moreover, research suggests that symptoms of ASD are identifiable as early as 12 months (Ozonoff et al., 2008; Rogers, 2009; Wetherby et al., 2004; Zwaigenbaum et al., 2005).

Typical toddlers utilize language to interact with others in the environment, leading to attachments and reciprocal relationships (e.g., Ainsworth & Bowlby, 1991; American Academy of Pediatrics, 2009; Rosen, Adamson & Bakeman, 1992; Striano & Rochat, 2000). Because toddlers develop these social skills at varying rates, atypicalities and delays in this area are more difficult to identify. However, research has identified specific social-emotional behaviors that are predictive of ASD. Lack of joint attention, limited use of gestures or pointing, absence of play initiation, poor language skills, limited response to name, and unconventional play with a limited variety of toys have all shown good predictive validity (Baranek, 1999; Gomez & Baird, 2005; Landa, Holman, & Garett-Mayer, 2007; Lord, 1995; Nadig et al., 2007; Osterling & Dawson, 1994; Wetherby et al., 2004). Based on these studies, delayed social communication skills are a major cause for concern in toddlers' development.

Typically developing toddlers display a number of challenging behaviors, such as hyperactivity, aggression, tantrums, defiance, and anxiety (Baillargeon et al., 2007; Bhatia et al., 1990; Osterman & Bjonkquist, 2010; Potegal & Davidson, 2003). Toddlers with ASD generally exhibit challenging behaviors more frequently and more intensely compared to typical children and children with other developmental disabilities (Matson, Mahan, Sipes, & Kozlowski, 2010; Matson, Wilkins, & Macken, 2009; Nicholas et al., 2008; Rojhan et al., 2009). For example, children with ASD show significantly more aggression (Dominick, Davis, Lainhart, Tager-Flusberg, & Folstein, 2007; Matson et al., 2009; Nicholas et al., 2008) and higher rates of attention difficulties (Hartley et al., 2008; Nicholas et al., 2008). Finally, many infants with ASD show regulation difficulties, with as much as 86% meeting criteria for a regulatory disorder at 1 year (Gomez & Baird, 2005). Richdale et al. have found that a significant percent of children with ASD experience sleep difficulties, such as problems falling asleep, staying asleep, shorter sleep cycles, and poor sleep quality (Cotton & Richdale, 2010; Patzold, Richdale, & Tonge, 1998; Richdale & Prior, 1995). In sum, the frequency and intensity of challenging behaviors in infants and toddlers are associated with later ASD diagnosis.

Additionally, some toddlers with ASD also display atypical behaviors (see Leekam, Prior, & Uljarevik, 2011 for a comprehensive review). Evaluation of repetitive, stereotyped behaviors of children under 2 years suggests the following behaviors reliably discriminate children with ASD from children with other developmental disorders and typical children: (1) repetitive hand, arm, or body movements; (2) placing hands to ears; (3) unusual posturing; (4) excessive mouthing of objects; and (5) aversion to social touching (Baraneck, 1999; Loh et al., 2007; Titelbaum, Teitelbaum, Nye, Fryman, & Maurer, 1998; Wetherby et al., 2004). Furthermore, a matched-control study by Watt, Wetherby, Barber, and Morgan (2008) found that children with ASD show significant differences in the frequency and duration of repetitive, stereotyped behaviors and sensory behaviors compared to typically developing children or children with developmental delays. Notably, two studies suggested children with ASD do not differ from children with developmental delay or typically developing children in regards to repetitive behaviors; however, they were limited by the use of a single parent report measure at 20 months (Cox et al., 1999) and analysis of home videos (Werner & Dawson, 2005).

In summary, the development of social communication skills along with the presence of challenging behaviors are inherent to toddlerhood. Individual differences in development of these nascent skills make the identification of ASD symptoms more difficult. However, research suggests that particular social communication deficits along with challenging and atypical behaviors are indicative of ASD (e.g., Gomez & Baird, 2005; Matson et al., 2010; Watt et al., 2008). Furthermore, reliable diagnoses of ASD can be made at 2 years (Charman & Baird, 2002; Kim & Lord, 2012; Lord, 1995; Lord et al., 2006; Stone et al., 1999).

Consistent with the average age of ASD diagnosis, the majority of research regarding intervention effectiveness has focused on preschoolers. Comprehensive reviews of early interventions showed significant improvements in cognitive, language, and adaptive functioning of children with ASD (Peters-Scheffer, Didden, Korzilius, & Sturney, 2011; Rogers & Vismara, 2008). The recent advancements in accurate diagnosis of toddlers has led to interest in developing comprehensive intervention programs for this population. Descriptive and quasi-experimental studies suggest promising results (Woods & Wetherby, 2003; Zwaigenbaum et al., 2009). An efficacy study of the Early Start Denver Model (Dawson et al., 2010) revealed improvements in cognitive, language, and adaptive skills along with a decrease in the severity of ASD symptoms as compared to a typical treatment plan received by toddlers with ASD. The increasing efficacy and availability of interventions for toddlers with ASD suggests that earlier identification of children with the disorder is warranted to ensure receipt of these services.

Early Screening

Universal screening practices can help identify children at risk for ASD by age two, thereby facilitating early diagnosis and intervention services. Various parent and medical groups advocate for the use of increased screening practices for ASD. The majority of young children have a primary care provider, usually a pediatrician, who monitors their development. Moreover, nine well-child visits are recommended by the child's second birthday (American Academy of Pediatrics, 2009). Specifically, standardized developmental screenings are recommended for all children at 9, 18, and 30 (or 24) months and targeted screenings for ASD at both 18 and 24 months (Johnson & Meyers, 2007). The next recommended well-child visit does not occur until 3 years so improved screening practices by 24 months is crucial to earlier identification of ASD.

One of the most significant barriers for ASD screening of toddlers is the limited availability of psychometrically sound assessments (Matson et al., 2008). The AAP (Johnson & Meyers, 2007) identified the M-CHAT (Robins, Fein, & Barton, 1999) as one of six recommended targeted screeners for use at the 18 month visit. Research on the M-CHAT suggests it has high sensitivity and specificity for ASD identification in children at least 18 months of age (Dumont-Matthieu & Fein, 2005; Kleinman et al., 2008, Pandey et al., 2008; Robins, Fein, & Barton, 1999).

Although many studies of ASD screeners are conducted with clinical samples; the M-CHAT has shown good reliability and validity in combined clinical and community samples (Dumont-Matthieu & Fein, 2005; Kleinman et al., 2008). The M-CHAT also effectively discriminated between ASD and non-ASD children when used by pediatricians at 2-year-old routine check-ups (Robins, Fein, Barton, & Green, 2001). Notably, a study by Pandey et al. (2008) found that 98% of children who screened positive on the M-CHAT at a well-child visit met diagnostic criteria for either ASD, language delay, or global developmental delay. More specifically, 79% met criteria for ASD. These findings suggest that the M-CHAT effectively screens for toddlers who need further evaluation for ASD and other developmental delays.

Although the M-CHAT may be the most widely used targeted ASD screener at the 24 month visit, it is not well supported for use in younger toddlers, especially before 18 months (e.g., Pandey et al., 2008, Robins & Dumont-Mathieu, 2006). Considering that most pediatricians conduct the general developmental screening as recommended (Dosreis, Weiner, Johnson, & Newschaffer, 2006), targeted ASD screening practices may be improved if a single instrument could serve both purposes. Moreover, identification of a brief screener that can be used across multiple toddler well-visits (i.e., 12, 15, 18, and 24 months) may also facilitate its use by pediatricians.

The Brief Infant and Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2006) is a parent report measure of social-emotional development and problem behaviors used for screening toddlers between 12 and 36 months. It is comprised of a social-emotional competence scale and a problem behavior scale with cut-points to identify children at risk for developmental problems. In addition to the norm-referenced scales, the BITSEA authors identify sets of content related items (i.e., externalizing, internalizing, dysregulation, ASD, and red flag). These item sets are derived from scales of the longer Infant Toddler Social Emotional Assessment; however, norms for these BITSEA subscales are not available (Briggs-Gowan & Carter, 2006). Considerable research by the authors reveals adequate to high reliability and validity in a birth cohort (Briggs-Gowan et al., 2004), national standardization (Briggs-Gowan & Carter, 2006), and early intervention (Briggs-Gowan & Carter, 2007) samples. Specifically, the BITSEA

Competence scale demonstrated excellent sensitivity (100%) and specificity (97%) identifying autistic disorder in 18 to 35 month old toddlers using (Briggs-Gowan & Carter, 2006). Notably, BITSEA psychometric properties have primarily been examined in largely white samples with less than 25% African American inclusion (Briggs-Gowan et al., 2006; Briggs-Gowan et al., 2001).

The validity of the BITSEA as a targeted screener for ASD was examined in a Turkish sample (Karabekiroglu et al., 2010). This study of the Turkish version of the BITSEA found that the Problem scale significantly correlated with the total score of the Autism Behavior Checklist ([AuBC]; Krug, Arick, & Almond, 1980) and the Competence score was significantly inversely correlated with the AuBC; indicating good convergent and discriminative validity of the scales as related to ASD (Karabekiroglu et al., 2010). Moreover, Competence scores in the autism group were found to be significantly lower when compared to typically developing children and those with disruptive behavior and internalizing disorders (i.e., anxiety and depression). These findings suggest that the BITSEA is a reliable and valid measure for identifying young toddlers at risk for developmental disorders and for discriminating between ASD and other developmental disorders in a Turkish sample.

The utility of the BITSEA as an ASD screener was also recently evaluated in a diverse (i.e., 65.4% African American, 32.5% Caucasian, 2.2% other) community sample of 1-year-olds (Gardner et al., 2013). First, Gardner et al. created a score for the ASD subscale comprised of the items identified by the BITSEA authors as related to ASD symptomology. Norms for the ASD scale were not provided by the authors (Briggs-Gowan et al., 2004), and this is the first study known to evaluate the predictive validity of these items for ASD risk. The ASD Total was compared to the Competence, Problem

Behavior, and Red Flag scales as predictors for ASD risk at 24 months as measured by the M-CHAT. This study found that the ASD Total was the best predictor of ASD risk at 24 months compared to the other scales. Although adequate sensitivity was obtained at 24 months (i.e., .76), it demonstrated low sensitivity (i.e., below .70) in 12 month olds. These results suggest further research is needed to improve the predictive validity of the scale in diverse samples.

Maternal Risk Factors

One method for improving classification accuracy is to include additional factors known to be associated with the outcome (Tabachnik & Fidell, 2013). Considerable research has identified a variety of maternal variables including maternal age as a risk factor for ASD (e.g., Bilder et al., 2009; Mandell et al., 2009; Shelton, Tancredi, & Hertz-Picciotto, 2010). In the most impressive of these studies, Shelton et al. (2010) examined maternal age and ASD risk by reviewing data from all California births from 1990 to 1999. Advancing maternal age was associated with increased risk for ASD, regardless of paternal age (except when mothers were less than 25 years old) while controlling for parent education, race/ethnicity, and socioeconomic status (based on insurance type). Similarly, Croen and colleagues (Croen, Grether, & Selvin, 2002; Croen, Najjar, Fireman, & Grether, 2007) identified both advancing maternal age (i.e., over 35 years) and higher levels of education (i.e., postgraduate) as risk factors for ASD, even when controlling for other demographic variables such as sex, birth weight, and ethnicity. Furthermore, a study by Mandell et al. (2009) reviewed records in 14 states and found that higher maternal education was associated with ASD. It is possible that parents with higher educational attainment may be more aware of ASD and therefore, more likely to pursue evaluation and services for their children.

Maternal mental health is an especially salient in predictor of a variety of social-emotional outcomes for children (e.g., Blandon, Calkins, & Keane, 2010; Carter et al., 2001; Cummings & Davies, 1994). A cohort study of all live births in Nova Scotia from 1990 to 2002 by Dodds et al. (2011) comprehensively evaluated risk factors for ASD. A history of maternal psychiatric illness was associated with a threefold likelihood of an ASD diagnosis. Furthermore, multiple studies have found higher rates of depressive symptoms in mothers of children with ASD compared to mothers of children with other disabilities and typical development (e.g., Davis & Carter, 2008; Dumas, Wolf, Fisher, & Culligan, 1991; Olsoon & Hwang, 2001). In sum, awareness of maternal psychological functioning may facilitate identification of toddlers' ASD risk.

Parenting stress is closely related to psychological distress, and is one of the most prominent parent characteristic in ASD research. Considerable research demonstrates high levels of parenting stress in families of preschool children with ASD (e.g., Koegel, Koegel, & Surratt, 1992; Sanders & Morgan, 1997; Smith, Oliver, & Innocenti, 2001) and in families with toddlers with ASD (Davis & Carter, 2008; Hastings, 2003; Tomanik, Harris, & Hawkins, 2004). Parenting stress increases with the severity of the diagnoses (Hastings & Johnson, 2001). More specifically, delays in social skills were most stressful for parents while cognitive and verbal delays were not predictive of parenting stress levels (Davis & Carter, 2008). On the other hand, a number of studies have shown that the presence of challenging behavior was more predictive of parenting stress than delays in competence (Baker, Blacher, Crnic, & Edelbrock., 2002; Baker et al., 2003; Nachshen, Garcin, & Minnes, 2005). Although the source of the stress is still under debate, together these findings suggest that mothers of toddlers with ASD are likely to display increased levels of parenting stress.

Purpose of this Study

Symptoms of ASD are evident as early as 12 months and accurate diagnosis can occur by the time a child is 2 years old. Moreover, effective interventions are available for toddlers and preschoolers with ASD (Peters-Scheffer et al., 2011, Zwaigenbaum et al., 2009). However, the average age of ASD diagnosis still remains at 3-4 years of age, suggesting that many children are not gaining access to the earliest intervention services available. Even more alarming is that African American children from lower income homes are at increased risk for delayed diagnosis (Liptak et al., 2008, Mandell et al., 2009), and they are more likely to receive a misdiagnosis of a behavior disorder (Mandell, 2007). Therefore, a need exists for earlier identification of minority children at risk for ASD, which can be accomplished through universal screening practices at the 18 and 24 month well child visits.

Emerging research suggests that screenings for ASD are both feasible and effective for identifying young toddlers at risk for ASD (Pierce, Carter, Weinfield, & Desmond, 2011). Unfortunately, only a small number of instruments are currently available for evaluating this risk (Matson et al., 2008). Effective screeners should be brief, inexpensive, and easy to administer and score. They also need to demonstrate good reliability and validity in community populations. Moreover, it is important for ASD screening instruments to demonstrate effectiveness with an African American population due to the disparity in diagnosis. Considering that pediatricians are more likely to conduct generalized developmental screenings, ASD screening practices may be improved if a single instrument can provide information on both general development (i.e., social-emotional development) and ASD risk. One such instrument is the BITSEA, which has shown promise as a developmental and targeted ASD screener in community samples

(Briggs-Gowan et al., 2006; Briggs-Gowan et al., 2004; Briggs-Gowan et al., 2001; Gardner et al., 2013; Karabekiroglu et al., 2010).

The psychometric properties of the BITSEA have primarily been examined in predominantly white samples (Briggs-Gowan et al., 2006; Briggs-Gowan et al., 2001), and its utility as an ASD screener has only begun to be explored. Recently, Gardner et al. (2013) examined the validity of the BITSEA as an ASD screener for 12- and 24-month-old toddlers with a diverse community sample drawn from the CANDLE study. The first purpose of the current study was to extend this research by focusing exclusively on the African American subgroup of longitudinal participants from the CANDLE study, including more than 450 additional African American participants for whom data was collected since the study conducted by Gardner et al. (2013).

The current study evaluated the following BITSEA scales: Competence, Problem, ASD, Dysregulation, and ASD + Dysregulation. In contrast to the study conducted by Gardner et al. (2013), the ASD scale was calculated by reverse scoring the competence items (i.e., 0 becomes 2 and 2 becomes 0) and adding it to the problem items. This approach eliminated the use of negative numbers in the scoring by creating a range of scores from 0 to 34 that would be easier for clinical interpretation. Additionally, the Dysregulation scale was examined, as research suggests that 12 month olds with ASD often have symptoms of dysregulation (e.g., Cotton & Richdale, 2010; Gomez & Baird, 2005). Risk for ASD was identified by total scores on the M-CHAT at year 2 (i.e., 0-2 = low risk and 3 or higher = at risk). Correlational and receiver operating characteristics (ROC) analyses were conducted for the BITSEA scales at year 1 (i.e., 12 months) and year 2 (i.e., 24 months) to identify the best ASD screening scale for both ages. Once identified, cut-points and classification statistics were also calculated for the most accurate ASD risk scale at each time point.

Although Gardner et al. (2013) found that the ASD scale was a better predictor of ASD risk than overall social emotional competence (as measured by the Competence and Problem Behavior scales), it produced inadequate sensitivity when used with 12-month-old children. The second purpose of the present research was to determine whether the classification accuracy of the BITSEA could be enhanced by including maternal factors associated with ASD. Specifically, age, education, income, parenting stress, and psychological distress were explored. Logistic regression models were computed to examine the contribution of maternal stress factors in predicting M-CHAT ASD risk status at year 2 after controlling for socio-demographic characteristics (i.e., age, education, and income). Separate models were created for 12-month-old and 24-month-old toddlers. Ultimately, the goal of this study is to inform screening practices in an effort to increase identification of African American toddlers at risk for ASD and facilitate early diagnosis and provision of intervention services.

Method

Participants

Participants were selected from the CANDLE study, which is an ongoing study evaluating prenatal and postnatal factors affecting early childhood development in an urban setting. Pregnant mothers were recruited using a variety of strategies, including announcements in local newspapers, hospitals, clinics, and other community agencies. The CANDLE study includes eight clinic visits occurring at various times from pregnancy until the child is 3 years old. Monetary compensation was provided at each of the visits, with a total of $500.00 available for full participation.

The current study includes 773 dyads of African American toddlers (51% boys) and their mothers with data from the clinic visits that occurred when the child was approximately 12 months old (i.e., year 1) and 24 months old (i.e., year 2). Only participants with child data from year 1 or year 2 were included in the current study (i.e., participants with only maternal data or data before year 1 were excluded). Toddlers were within 6 months of the targeted age (i.e., 12 and 24 months) at time of testing. Notably, 1 participant was excluded from the study because testing occurred outside of this targeted age range at the first clinic visit and no other data were available. The average age of children in the final sample was 12.64 months (*SD* = 1.64) at year 1 and 24.48 months (*SD* = 1.46) at year 2. Notably, 8.4% of children were born preterm (i.e., before 37 weeks); however, the average gestational age for the sample is 38.73 weeks.

According to self-report data, the average age of mothers was 26.58 years (*SD* = 5.37) with the majority (70%) between 21 and 31 years. Approximately half of the mothers reported a high school diploma or GED as their highest level of education (55.9%) and were never married (55%). Most of the participating children (72.2%) were uninsured or on Medicaid insurance, according to maternal report. Refer to Table 1 for a summary of demographic characteristics by M-CHAT ASD risk status at year 2.

Measures

A range of assessments was conducted in the CANDLE study; however, only data from the following measures were used in the current study: BITSEA; M-CHAT; Parenting Stress Index, Short Form (PSI; Abidin, 1995b); and Behavior Symptoms Inventory (BSI; Deragatis, 1993). The BITSEA, PSI, and BSI were administered at both year 1 and year 2 clinic visits, while the M-CHAT was administered only at year 2. Raw score data from these evaluations were entered by the examiner (i.e., BITSEA and M-

Table 1

Participant Characteristics at Time 1

	Total (N=773)		Year 2 M-CHAT *low risk*[a] *(n=717)*		Year 2 M-CHAT *moderate to high risk*[b] *(n=56)*	
	N	*%*	*n*	*%*	*n*	*%*
Child's Gender						
Boy	391	50.6	356	49.7	35	62.5
Girl	382	49.4	361	50.3	21	37.5
Race of Mother						
African American	765	99.0	709	98.9	56	100
White	7	0.9	7	1	0	0
Other	1	0.1	1	0.1	0	0
Missing	0	0	0	0	0	0
Marital Status						
Never married	425	55.0	388	54.1	37	66.1
Married	174	22.5	164	22.9	10	17.9
Living with partner	139	18.0	130	18.1	9	16.1
Divorced	17	2.2	17	2.4	0	0
Separated	13	1.7	13	1.8	0	0
Widowed	1	0.1	1	0.1	0	0
Missing	4	0	4	0.6	0	0
Educational Status						
High school diploma or GED	432	55.9	392	54.7	40	71.4
College degree	132	17.1	128	17.9	4	7.1
Technical school	95	12.3	92	12.8	3	5.4
Graduate/professional degree	54	7.0	52	7.3	2	3.6
Less than high school	55	7.1	48	6.7	7	12.5
Missing	5	0.6	5	0.7	0	0
Insurance Status						
Medicaid	490	63.4	447	62.3	43	76.8
Other	208	26.9%	201	28.0%	7	12.5%
None	68	8.8%	62	8.6%	6	10.7%
Missing	7	0.9%	7	1.0%	0	0

CHAT) and research assistants (i.e., PSI and BSI) into an online data management system. The online system calculated standard scores used in the analyses utilizing the normative data from the respective test manuals.

Autism spectrum disorder risk. Two instruments were used to assess ASD risk in this study: the BITSEA (Briggs-Gowan & Carter, 2006) and the M-CHAT (Robins, Fein, & Barton, 1999). The BITSEA is being evaluated in this study to determine its usefulness as an ASD screener for children aged 1 year and 2 years. The M-CHAT is a widely used screener for children 16 months to 30 months and is utilized in this study as the outcome indicator of ASD risk status at year 2.

BITSEA. The BITSEA is a parent report of social-emotional problem behaviors and developmental competencies for children ages 12 months to 35 months 30 days. The parent may complete the form independently as a questionnaire or it may be administered as a brief interview, as utilized in this study. Responses are rated on a 3-point scale with 0 = not true/rarely, 1 = somewhat true/sometimes, and 2 = very true/often. The BITSEA yields Problem Behavior and Competence raw scores and uses cut points to identify children at risk for social-emotional and behavioral problems; percentile scores are also available. The Problem scale consists of 31 items addressing internalizing, externalizing, and atypical behaviors such as anxiety, aggression, and repetitive behaviors. The Competence scale is comprised of 11 items addressing prosocial skills such as empathy, compliance, and secure attachment. High Problem scores indicate the presence of problem behaviors, whereas low Competence scores indicate possible deficits in social skills.

Although a number of studies have examined the psychometric properties of the BITSEA, the results of a study in a community birth cohort sample (Briggs-Gowan et al.,

2004) are provided here because the current sample is also a community sample. Test–retest reliability coefficients for the BITSEA Problem scale and Competence scales were excellent at .87 and .85, respectively. Moreover, one-year stability was r = .65 for Problems and r = .53 for Competence. Inter-rater reliability coefficients were .68 and .61 for the Problem and Competence scales, respectively. The BITSEA Problem and Competence scores were found to be moderately to strongly correlated (i.e., .45 to .63) with the associated Infant-Toddler Social and Emotional Assessment (ITSEA) domains. Finally, at the 10th percentile cutpoint, the BITSEA competence scale demonstrated excellent sensitivity (100%) and specificity (97%) predicting autistic disorder (Briggs-Gowan & Carter, 2006).

In addition to the norm-referenced scales, the BITSEA authors identify subsets of items related to more specific problem areas; however, norms for these scales are not available (Briggs-Gowan & Carter, 2006). Specifically, 17 items related to ASD symptomology (e.g., empathy, social referencing, atypical behaviors) are indicated across both the Competence and Problem scales. The authors suggest that problems and deficits on these items may indicate the need for follow-up, regardless of scores on the overall Competence and Problem behavior scales. Although the items are identified in the manual, no scoring criteria are suggested. As this study seeks to examine the utility of these items, an ASD score was created. Competence items are reversed scored (i.e., a score of 0 will become a score of 2) and added to the problem items so that higher scores indicate more symptomology.

Additionally, items related to dysregulation (i.e., negative emotionality, sleeping difficulties, eating problems, and sensory sensitivities) are identified by the BITSEA authors, which may suggest the need for follow-up. Considering that many children with

ASD also show signs of dysregulation at 12 months (e.g., Cotton & Richdale, 2010; Gomez & Baird, 2005), this scale may help identify toddlers at risk for ASD, although it has not yet been examined for that purpose. All 8 dysregulation items are problem behaviors; therefore, scores across these items were summed to create the Dysregulation score. A combined ASD + Dysregulation scale was also created for use in this study by summing the two scales.

M-CHAT. The M-CHAT is a screener for ASD with a pleathora of studies supporting its utility for this purpose. The M-CHAT is intended for use with children 16 to 30 months of age. It is comprised of 23 yes/no items to be completed by parent/caregiver directly or through interview. Total raw scores of 3 or higher suggest moderate to high risk for ASD. The M-CHAT is an American adaptation of the Checklist for Autism in Toddlers, which was developed in the United Kingdom. Follow-up questions are suggested based on parent responses, which significantly improves the validity of the instrument (Robins et al., 2001). In this study, these questions were administered at the time of administration, when indicated.

The M-CHAT norming sample is based on a sample of 1,293 toddlers. Post hoc discriminate function analysis showed significant differences between ASD and non-ASD toddlers on M-CHAT scores. Internal reliability was judged adequate with Cronbach's alpha of .83 (Robins & Dumont-Mathieu, 2006). Kleinman et al. (2008) found that the M-CHAT with the follow-up interview accurately flagged at risk 74% of children who were later diagnosed with ASD. This finding is consistent with previous literature demonstrating that the M-CHAT adequately predicts ASD diagnosis (Robins et al., 2001). Additionally, Robins (2008) found M-CHAT sensitivity at .87 and specificity at .99 in predicting ASD; further supporting the utility of the instrument as an ASD screener. Two studies also have examined risk for ASD in high-risk preterm infants using the M-CHAT as their outcome variable; however, both acknowledge that it serves only as a screener and that additional diagnostic evaluations were warranted (Kuban et al., 2009; Limperopoulos et al., 2008). Compared to other available screeners for toddlers, the M-CHAT has wide research base of psychometric studies demonstrating its validity and reliability as an ASD screener. Therefore, the M-CHAT total score is used in this study to identify children at risk for ASD at year 2 (henceforth referred to as M-CHAT risk).

Maternal stress. Two measures of maternal stress were also included in this study to examine if identification of parenting stress and psychological distress could facilitate predication of ASD risk at either year 1 or year 2. The Parenting Stress Index-Short Form (PSI; Abidin, 1995b) is a parent questionnaire that evaluates areas of parent-child relationships that are problems. The PSI takes approximately 15 min. to complete and contains 36 items taken from the longer version (Abidin, 1995b). Ratings are measured on a 5-point scale ranging from 1=*strongly agree* to 5=*strongly disagree*. A total stress composite score is obtained from 3 scales (i.e., difficult child, parent-child dysfunctional interaction, and parental distress). The PSI-SF manual reported excellent internal consistency (Cronbach's a = .91) for Total Stress. Preliminary analyses revealed that the Total Stress score was significantly correlated with the M-CHAT scores (year 1 r =.15, p < .001; year 2 r = .31, p < .001). This score was used in subsequent analyses to represent overall parenting stress. According to the manual, scores greater than 90 are considered clinically significant (M = 50, SD = 10).

The Brief Symptoms Inventory (BSI; Derogatis, 1993a) is a 53-item parent questionnaire adapted as a brief form of the Symptoms Checklist-90-Revised, with an estimated administration time of 15 minutes (Derogatis, 1993b). Ratings of “how much that problem has distressed or bothered you during the past 7 days including today" are based on a 5-point scale ranging from 0 = *not at all* to 4 = *extremel*y (Derogatis, 1993b). The BSI assesses psychological symptom patterns across nine primary dimensions, including Interpersonal Sensitivity, Depression and Anxiety. Three composites are also produced to provide overall indicators of psychological well-being (i.e, Global Severity Index [GSI], Positive Symptoms Total, and Positive Symptoms Distress Index). The GSI is widely accepted as an overall measure of psychological distress (see Skeem et al., 2006). Moreover, preliminary analyses revealed that the GSI was more highly correlated with the M-CHAT scores at both year 1 and year 2 ($r = .14$, $p < .001$ and $r = .18$, $p < .001$, respectively) compared to the other composites. Hence, the GSI was used as an indicator of overall psychological distress in the current study. According to the manual, GSI T scores greater than or equal to 63 are considered clinically significant ($M = 50$, $SD = 10$).

Procedures

Mothers provided informed written consent for participation in the study at enrollment and all subsequent visits. Demographic information (e.g., maternal age and education level) were provided at the enrollment visit and updated at each subsequent visit. All other data for the current study are drawn from two of the clinic visits; therefore, only those procedures are described. Visits were scheduled in the mornings and early afternoons and lasted approximately 3.5 hr. overall, including measures not used in the present study. Psychological examiners who conducted the assessments received extensive training from licensed psychologists at the Boling Center for Developmental Disabilities at the University of Tennessee Health Sciences Center in Memphis, Tennessee. The small group training occurred over 2 days and involved direct instruction of assessment procedures and in-vivo practice administrations. Following the didactic training sessions, inter-rater reliability sessions were scheduled; all cognitive examiners met or exceeded 90% agreement with a senior examiner.

Data Analysis

One purpose of this work was to replicate and extend the Gardner et al. (2013) study examining the psychometric properties of the BITSEA with an African American sample. To this end, correlational and ROC analyses were conducted with the BITSEA scales from year 1 and year 2, using the year 2 M-CHAT scores as the criterion measure for ASD risk. The second purpose of this study was to explore the potential contribution of maternal factors in identifying ASD risk. Logistic regression models were used to examine the contribution of maternal factors, in combination with the BITSEA scores at year 1 and year 2, to predict ASD risk status as measured by the year 2 M-CHAT assessment.

Correlations. Correlational analyses were completed on the following five BITSEA scales: (1) Competence, (2) Problems, (3) Dysregulation, (4) ASD, and (5) ASD + Dysregulation. To examine temporal stability, Pearson's product-moment correlations were calculated for the year 1 and year 2 scores on the five BITSEA scales. Although test–retest reliability of .70 is generally considered acceptable (Nunnally, 1978), it is important to note that the time between tests was approximately 1 year and toddler's social-emotional skills develop at varying rates at this age (Briggs-Gowen & Carter, 1998; Mathieson & Sanson, 2000; Mesman & Koot, 2001). Internal consistency for each of these scales also was calculated separately at each year using Cronbach's alpha. The following guidelines were used to describe internal consistency in this study: *poor,* .5 to .59; *acceptable,* .6 to .69; *good,* .7 to .89; excellent, .9 or higher (Kline, 2000).

Additionally, predictive and concurrent validity of the year 1 and year 2 BITSEA scales were examined by computing Pearson's product-moment correlations for the 5 scales and the year 2 M-CHAT total score. The following guidelines are used to describe these correlational results: *negligible,* .00 to .19; *weak,* .20 to 39; *moderate,* .40 to .69; *strong,* .70 to .89; and *very strong,* .90 to 1.0 (Floyd et al., 2008). Alpha was set at .01 to be identified as statistically significant.

ROC analyses. The classification accuracy of the 5 BITSEA scales was also examined through receiver operator characteristics (ROC) curve analyses. The area under the curve (AUC) provides a measure of test accuracy by indicating how well a test classifies cases into the correct group. This study used the dichotomous outcome variable of M-CHAT risk status, where scores of 0-2 indicate low risk and scores of 3 or higher indicate moderate to high risk. Separate analyses were completed for year 1 and year 2 BITSEA scales. An AUC value of .50 indicates that a measure has chance accuracy in predicting an outcome. AUCs falling between .50 and .70 indicate low test accuracy, .70 to .90 indicate moderate test accuracy, and .90 to 1.0 indicate high test accuracy (Swets, 1998).

Next, results from the correlational and AUC analyses were used to identify the most predictive scales, and these selected scales were used in subsequent ROC curve analyses. Classification statistics were calculated for the most predictive BITSA scales at each year to identify possible cutpoints. True positives are the number of children accurately identified as at risk on the BITSEA that also scored at risk (i.e., moderate to high risk) on the M-CHAT. False negative refers to the children who were not identified at risk on the BITSEA; however, they scored at moderate to high risk on the M-CHAT. The ROC curve plots the true positive rate (i.e., sensitivity) against the false positive rate (i.e., 1-specificity) across the range of all possible scores for the scale. This ROC curve analysis allowed for identification of the cutpoints that yield various sensitivity and specificity levels. Sensitivity and specificity of at least 80% is preferred (Glascoe, 1991). Notably, sensitivity may be prioritized in screeners because too many false negatives would deny evaluation and intervention services to children in need. Considering that practitioners may prefer different levels of sensitivity, classification statistics for multiple cutpoints for each year were calculated using cross tabs. Optimal cutpoints with approximately 80% sensitivity at year 1 and year 2 were highlighted.

Logistic regression. Logistic regression analyses were employed to examine the contribution of maternal variables in identifying ASD risk in toddlers as measured by the M-CHAT. Only 56 participants (7.2%) met ASD risk criteria (i.e., total scores of 3 or higher) on the M-CHAT. To maximize power and reduce bias, it is recommended to limit the ratio of predictor variables to positive events (i.e., at risk for ASD) to about 1:10 (Peduzzi, Concato, Feinstein, & Holford, 1995). Therefore, the number of predictor variables in the logistic regression was limited to 5 or less for each year. Initially, univariable analyses were conducted with each potential risk factor (i.e., demographic variables, maternal stress variables, and the BITSEA ASD scales) to identify significant factors ($p < .25$) to include in the model. The following factors were significant: maternal education, insurance status, parenting stress, psychological distress, and BITSEA ASD scale score. Maternal age was also included in the initial model due to the plethora of research suggesting its importance.

Separate logistic regression models were then computed for year 1 and year 2 BITSEA scores, and analyses were conducted using a three-step process. In step 1, the demographic variables (i.e., maternal education, health insurance status, and maternal age) were entered blockwise as potential within-time predictors in the logistic model. Then in step 2, the selected BITSEA ASD score (i.e., year 1 = ASD + Dysregulation score, year 2 = ASD score) was entered into the model. Finally, in step 3, the maternal stress variables (i.e., psychological distress and parenting stress) were entered into the model. Multivariate models were developed using backward elimination regression to test the significance of each factor. Variables that were significant (i.e., $p < .10$) based on the Wald test were included in the final model. Additionally, highly correlated variables were examined and reduced when appropriate to limit multicollinarity. Changes in log likelihood were also examined to determine if the addition of maternal stress variables contributed significantly to the model. Wald statistics, significance levels, odds ratios and confidence intervals were calculated for the variables in the final models.

Results

Data Screening and Handling of Missing Data

As noted, data screening revealed that one child had been tested outside the appropriate testing window for year 1 (i.e., child older than 18 months). This child did not have any other data available; therefore, he was excluded from the database. Two other children were tested outside the appropriate testing window at year 2 (i.e., older than 30 months); however, they had valid data from year 1 that was included in the study. Data were also screened for outliers, distributional properties, and parametric assumptions (Tabachnik & Fidell, 2013). Although some outliers were identified (i.e., $z > 3.29$; Tabachnick & Fidell, 2013), deleting them suggested that they were not influential data points and therefore, they remained unchanged in the data set.

Approximately 10% of participants had missing data at year 1. Approximately 15% of participants did not return for the year 2 clinic visit and were missing all year 2 data. Furthermore, 22% of participants had complete child data but were missing maternal data for year 2. The missing data from these participants were examined using Little's Missing Completely at Random test (Little, 1988). Results yielded a statistically significant result, suggesting that the missing values were not missing completely at random. The SPSS expectation-maximization algorithm was used to impute the missing data for these cases. Results were comparable with and without the imputed data; therefore, the complete data with missing data imputed was used in these analyses.

Correlations

Temporal stability was examined by computing correlations between the year 1 and year 2 scores on the following BITSEA scales: (1) Competence, (2) Problems, (3) Dysregulation, (4) ASD, and (5) ASD + Dysregulation. One-year stability was significant ($p < .001$); however, weak correlations were found for all the scales. Internal consistency also was examined for each of these scales and results were comparable across year 1 and year 2, with Cronbach's alphas ranged from poor to good (.54 to .79). Table 2 provides a complete description of these reliability statistics for the BITSEA scales.

Table 2

Reliability of the BITSEA scales

BITSEA scale	No. of items	Stability Year 1 to Year 2	Internal consistency Year 1	Internal consistency Year 2
Competence	11	.28*	.60	.69
Problems	31	.39*	.77	.79
ASD	17	.31*	.54	.67
Dysregulation	8	.30*	.56	.55
ASD + Dysregulation	25	.35*	.58	.56

Note. $*p < .001$.

The year 1 and year 2 BITSEA scales were also correlated with the M-CHAT total score (see Table 3). All correlations were significant and in expected directions, with stronger relations observed in year 2. Correlations between the BITSEA scales and the M-CHAT total score ranged from negligible to weak at year 1, whereas values ranged from weak to moderate in year 2. Notably, the BITSEA ASD scale and the BITSEA ASD + Dysregulation scale had the highest correlations with the M-CHAT total scores across both time points. However, the ASD + Dysregulation scale demonstrated the strongest relation to the M-CHAT total score ($r = .23$, $p < .001$) in year 1, and the ASD scale demonstrating the highest correlation with the M-CHAT total score ($r = .46$, $p < .01$).

Table 3

BITSEA Descriptive Statistics and Correlations with the M-CHAT total score

BITSEA scale	*M*	*SD*	*r*	AUC	*SE*
Year 1					
Competence	15.06	3.04	-.18*	.62	.04
Problem	10.99	5.86	.17*	.65	.04
ASD	6.36	3.12	.22*	.67	.03
Dysregulation	3.67	2.38	.15*	.62	.04
ASD+ Dysregulation	9.05	3.95	.23*	.68	.03
Year 2					
Competence	17.78	2.54	-.42*	.76	.04
Problem	11.04	5.78	.33*	.69	.04
ASD	3.71	3.01	.46*	.86	.02
Dysregulation	3.69	2.21	.22*	.66	.05
ASD + Dysregulation	7.39	4.34	.43*	.83	.03

Note. $N = 773$; Brief Infant-Toddler Social and Emotional Assessment (BITSEA); Problem = BITSEA Total Problem Raw Score; Competence = BITSEA Total Competence Raw Score; ASD = BITSEA ASD Scale Raw Score; Dysregulation = BITSEA Dysregulation Scale Raw Score; ASD + Dysregulation = sum of the BITSEA ASD and Dysregulation Scales. *$p < .001$.

ROC Analyses

Area under the curve (AUC) was also calculated for each of the scales using M-CHAT risk status (i.e., 0-2 = low risk and 3 or higher = moderate to high risk) as the outcome variable. Consistent with the correlational analyses, the ASD and ASD + Dysregulation scales had the highest AUC among the BITSEA scales and both produced moderate test accuracy at year 1 and year 2. Specifically, at year 1 the ASD + Dysregulation scale had a slightly higher AUC compared to the ASD scale alone (AUC = .683 and AUC = .668, respectively). At year 2, the ASD scale was higher with AUC = .860 compared to the combined scale with an AUC = .834. Based on these analyses, the most accurate BITSEA scales in identifying toddlers at risk for ASD were the ASD + Dysregulation at year 1 and ASD alone at year 2. Therefore, these scales were used in subsequent analyses to represent BITSEA ASD risk.

Next, ROC curves were evaluated to determine possible cutpoints for the BITSEA ASD risk score at both year 1 and year 2. Generally, an optimal cutpoint maximizes sensitivity and specificity; however, sensitivity may be prioritized in this screener (i.e., 80% to 90%). The M-CHAT risk status again was used as the outcome variable in these analyses. At year 1, a cutpoint of 9 on the ASD + Dysregulation scale produced sensitivity at 79.6% and specificity at 49.2%. At year 2, a cutpoint of 6 resulted in sensitivity at 80.4% and specificity at 81.2%. Refer to Table 4 for a summary of classification statistics for various cutpoints.

Logistic Regression

Logistic regression analyses were employed to examine the contribution of maternal factors along with the BITSEA in predicting whether participants were at risk for ASD as measured by the M-CHAT. Separate models were computed using the

Table 4

Classification Statistics Using BITSEA ASD to Identify M-CHAT ASD Risk at Year 2

BITSEA score	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
Year 1 (ASD + Dysregulation)				
Cutoff 7	92.9	27.2	9.1	98.0
Cutoff 8	87.5	36.6	9.7	97.4
Cutoff 9	79.6	49.2	10.6	97.0
Cutoff 10	58.9	63.0	11.1	95.2
Year 2 (ASD)				
Cutoff 4	89.3	61.6	15.4	98.7
Cutoff 5	82.1	72.5	18.9	98.1
Cutoff 6	80.4	81.2	25.0	98.1
Cutoff 7	66.1	88.0	30.1	97.1

Note. Cutoff score is defined as the lowest number indicating "positive" outcome or ASD risk. Sensitivity is calculated as the ratio of positives on the BITSEA to positives on the M-CHAT (i.e., true positives). Specificity is the ratio of negatives on the BITSEA to negatives on the M-CHAT (i.e., true negatives). The positive predictive value is the number of true positives compared to the total number of children identified as at risk on the BITSEA. Negative predictive value indicates the number of true negatives compared to the total number of children identified "low risk" on the M-CHAT.

selected BITSEA scales for each year (BITSEA ASD + Dysregulation in year 1; BITSEA ASD in year 2) with the M-CHAT ASD risk status at year 2 as the outcome variable. Preliminary univariable analyses and backward elimination stepwise procedures guided independent variable selection for the final models. Wilcoxon-Mann-Whitney tests indicated significant differences between the M-CHAT low risk and M-CHAT moderate to high risk groups on all predictor variables (i.e., BITSEA ASD risk, psychological distress, and parenting stress) at both years (see Table 5). All final variables were significantly related to M-CHAT risk at year 2 (see Table 6).

Table 5

Descriptive Statistics of Predictor Variables Grouped by Year and ASD Risk Status

	Year 2 M-CHAT low risk[a] (*n* =717)		Year 2 M-CHAT moderate to high risk[b] (*n* =56)	
Predictor Variables	*M*	*SD*	*M*	*SD*
Year 1				
BITSEA ASD+Dysregulation	8.87	3.91	11.36	3.64
Psychological Distress	46.94	9.93	52.95	11.29
Parenting Stress	60.47	17.82	71.38	18.57
Year 2				
BITSEA ASD	3.37	2.68	7.99	3.63
Psychological Distress	45.60	9.54	52.66	12.15
Parenting Stress	62.52	17.40	78.20	23.84

Note. Low risk and at risk groups were significantly different on all variables at $p < .001$ level. [a] Low risk status determined by M-CHAT scores of 2 or lower. [b] Moderate to high risk status determined by M-CHAT scores of 3 or higher.

Table 6

Descriptive Statistics and Correlations of the Predictor Variables and M-CHAT

Predictor Variables	*M*	*SD*	1	2	3	4	5	6
Year 1								
1 BITSEA ASD+Dysregulation	9.05	3.95	--					
2 Psychological Distress	47.37	10.14	.24*	--				
3 Parenting Stress	61.26	18.09	.36*	.52*	--			
Year 2								
4 BITSEA ASD	3.71	3.01	.31*	.12*	.20*	--		
5 Psychological Distress	46.11	9.91	.19*	.67*	.37*	.15*	--	
6 Parenting Stress	63.66	18.38	.27*	.37*	.59*	.25*	.47*	--
7 M-CHAT ASD total score	.82	1.18	.23*	.15*	.16*	.46*	.17*	.28*
8 M-CHAT ASD risk status	.07	.26	.16*	.13*	.16*	.40*	.19*	.22*

Note. $*p < .001$.

The final logistic regression models were constructed in a three-step process as described in the Method section. This procedure resulted in the same predictor variables for both year 1 and year 2. Maternal age, education, and health insurance status were initially entered in step 1. When health insurance was included in the model, maternal age and education were not significant and were eliminated from the model. Note that health insurance was identified as a categorical variable, which was defined as “no insurance/Medicaid” versus “private insurance.” In step 2, the BITSEA score was entered. Based on analyses described earlier in this paper, the ASD + Dysregulation scale was used in year 1 and the ASD scale was used in year 2. In step 3, the maternal variables, psychological distress and parenting stress, were entered into the model. Notably, these variables were moderately correlated to each other (year 1 $r = .52$; year 2 $r = .47$) and became nonsignificant when both were included. Therefore, the stronger variable, psychological distress, remained in the final model and parenting stress was eliminated.

Tests of the constant-only model compared to each step and the full models at both years yielded significant results, indicating that each predictor variable contributed significantly to the models. Both year 1 and year 2 final models passed the Hosmer and Lemeshow “lack-of-fit” tests, year 1 X^2 (8, N = 766) = 9.804; $p = .279$ and year 2 X^2 (8, N = 766) = 9.629; $p = .292$. Table 7 shows summary statistics for the final models.

In year 1, ROC curve analysis of the final model revealed an AUC of .870, suggesting good accuracy in predicting M-CHAT ASD risk. The odds ratios (OR) indicate that participants with private insurance were more than twice as likely to be at risk for ASD compared to those on Medicaid or without insurance (OR = 2.22). For a one point increase on the BITSEA ASD + Dysregulation scale the odds of M-CHAT ASD risk are multiplied by 13% (OR = 1.13). Furthermore, the addition of psychological distress added significantly to the model beyond demographic and ASD characteristics

Table 7

Logistic Regression Analysis of BITSEA and Maternal Variables on M-CHAT Risk

Variable	B	*SE*	Wald	*p*	Exp(B)	95% C.I. For Exp(B)
Year 1						
Insurance						
Private	.797	.419	3.610	.057	2.219	.975 – 5.049
BITSEA ASD+Dysregulation	.111	.033	11.047	.001	1.117	1.046 – 1.192
Psychological distress	.040	.014	8.240	.004	1.041	1.013 – 1.070
Constant	-6.304	.800	62.142	.000	.002	
Lack of fit x^2 =	9.804	$p = .279$				
Pseudo R^2 =	.104					
AUC =	.870	$p < .001$				
Year 2						
Insurance						
Private	.483	.443	1.185	.276	1.620	.679 – 3.865
BITSEA ASD	.373	.046	64.923	.000	1.453	1.327 – 1.591
Psychological distress	.056	.015	13.551	.000	1.058	1.027 – 1.090
Constant	-7.686	.916	70.420	.000	.000	
Lack of fit x^2 =	9.629	$p = .292$				
Pseudo R^2 =	.327					
AUC =	.867	$p < .001$				

Note. The dependent variable in this analysis is M-CHAT risk status coded so that 0 = low risk and 1 = moderate to high risk, *N* = 766.

alone ($p < .001$). A one point increase on the psychological distress scale increased the odds of ASD risk by 4% (O.R. = 1.04). Notably, the final formula at year 1 for predicting probability of M-CHAT ASD risk is

$$p = \frac{e^{-6.304 + .797\,(\text{insurance}) + .040\,(\text{psychological distress}) + .111\,(\text{ASD + Dysregulation})}}{1 + [\,e^{-6.304 + .797\,(\text{insurance}) + .040\,(\text{psychological distress}) + .111\,(\text{ASD + Dysregulation})}]}$$

For example, a 1-year-old without insurance or on Medicaid, whose mother indicates a significant level of psychological distress (i.e., BSI total score = 63), and who receives a score of 9 on the BITSEA ASD + Dysregulation scale has a 6% chance of being at moderate to high risk for ASD (i.e., M-CHAT risk). The same child with private insurance would have a 12% chance of scoring positively.

In year 2, ROC curve analysis of the model revealed an AUC of .867 ($p < .001$), suggesting good accuracy in predicting M-CHAT ASD risk. Participants with private insurance were about 1.5 times as likely to be at risk for ASD as those on Medicaid or without insurance (O.R. = 1.62). A one point increase on the BITSEA ASD scale increased the odds of M-CHAT ASD risk by 45% (O.R. = 1.45). The psychological distress O.R. indicated that when controlling for insurance and ASD symptoms, the psychological distress scale contributed significantly to the model ($p < .001$) and increased the odds of M-CHAT ASD risk by 6% (O.R. = 1.06). Notably, the final formula for predicting the year 2 probability of M-CHAT ASD risk is

$$p = \frac{e^{-7.686 + .483\,(\text{insurance}) + .056\,(\text{psychological distress}) + .373\,(\text{ASD})}}{1 + [\, e^{-7.686 + .483\,(\text{insurance}) + .056\,(\text{psychological distress}) + .373\,(\text{ASD})}]}$$

For example, a 2-year-old with no insurance or Medicaid, whose mother indicates a significant level of psychological distress (i.e., BSI total score = 63), and who receives a score of 4 on the BITSEA ASD scale has a 12.8% chance of being at risk for ASD (i.e., M-CHAT risk). The same child with private insurance has a 19.2% chance of a positive score.

Discussion

ASD can be reliably diagnosed as early as 18 to 24 months (Kim & Lord, 2012; Lord et al., 2006) and interventions for toddlers can be effective (Zwaigenbaum et al., 2009). Although early identification is possible, most children are not diagnosed until about 3 to 4 years of age (Biao, 2012; Howlin & Moore, 1997; Mandell et al., 2005).

Furthermore, African American children under 6 years of age have the lowest rate of diagnosis compared to other ethnicities (Liptak et al., 2008) and they are more likely to receive a misdiagnosis of a behavior disorder (Mandell, 2007). Universal screening practices, such as those recommended by the AAP, can facilitate these earlier diagnoses and increase access to early intervention services. The BITSEA has demonstrated strong psychometric properties as a social-emotional screener in diverse populations (Briggs-Gowan & Carter, 2004) and has also shown promise as a targeted screener for ASD (Gardner et al., 2013; Karabekiroglu et al., 2010). The results of this study add to this growing literature base by demonstrating the utility of the BITSEA as an ASD screener in an African American community sample.

Psychometric Properties of the BITSEA

Reliability analyses were restricted to one-year stability and internal consistency because these data were extrapolated from the larger, longitudinal CANDLE study. Research on the stability of traits from infancy to toddlerhood suggests correlations range from low to moderate for internalizing problems and moderate to strong for externalizing problems (e.g., Briggs-Gowen & Carter, 1998; Mathieson & Sanson, 2000; Mesman & Koot, 2001). Results from the current study indicated low stability of social-emotional and ASD traits from year 1 to year 2. Although the correlations were lower than desired, weak relationships were expected because the time between testing sessions was lengthy (i.e., one year). Internal consistency was also calculated for each of the scales in both years. Both the ASD and ASD + Dysregulation scales demonstrated poor internal consistency in year 1. On the other hand, internal consistency for the Competence and ASD scales improved in year 2 to adequate levels. This finding aligns with research that suggests nonverbal communication and ASD characteristics (particularly related to social competence and communication) begin to stabilize at around 2 to 3 years (Charman & Baird, 2002; Lord et al., 2006).

Current and previous research indicate that the BITSEA shows promise as an ASD screener. The BITSEA Problem and Competence scales have strong evidence to support their use as a screener for ASD in predominantly white samples (see Briggs-Gowan & Carter, 2006) and a Turkish sample (Karabekiroglu et al., 2010). Using the items identified by the BITSEA authors as related to ASD, Gardner et al. (2013) found that a total ASD score comprised of these items better indicated ASD risk compared to the Competence, Problem, Red Flag items, ASD Competence items, and ASD Problem items in a community sample. Consistent with the Gardner et al. (2013) findings, results from correlational and AUC analyses suggested the ASD scale outperforms the Problem and Competence scales as an indicator of ASD risk at both year 1 and year 2. Notably, correlational results indicated significant yet weak relations between the BITSEA scales and the M-CHAT total score. A possible explanation for this finding is the restricted range and limited variability in M-CHAT scores with most children scoring between 0 and 2 (i.e., low risk).

The combined ASD + Dysregulation scale at year 1 also significantly predicted ASD risk at year 2. As many children with ASD demonstrate difficulties with regulation at 12 months (Gomez & Baird, 2005), this finding was expected. Although both the correlation and AUC were slightly higher for the combined ASD and Dysregulation scale, the inclusion of dysregulation items did not appear to significantly improve the predictive validity of the ASD scale at year 1. Although beyond the scope of this study, it is possible that a combination of specific dysregulation items and the ASD scale may be important in the prediction of ASD.

The next step in making the BITSEA more usable as an ASD screener was to identify cutpoints for the ASD scales through examination of classification statistics. Due to the low base rate of ASD and use of the BITSEA scale as a screener, higher sensitivity with over identification of at risk children is desired and lower specificity is expected (Baldessarini, Finkelstein, & Arana, 1983; Canyon & Goodstein, 1982; Cronbach, 1984). The Gardner et al. (2013) study is the only previous research known to evaluate the BITSEA ASD scale. The method used by authors to calculate an ASD score resulted in negative numbers; therefore, the current study reverse scored competence items to yield only positive scores for easier interpretation. This change resulted in an inability to directly compare the cutpoints in the two studies; however, classification statistics can be examined.

Neither study could identify a cutpoint for the ASD total scale in detecting 1-year-olds at risk for ASD that produced sensitivity and specificity above .70. The current study found that a cutpoint of 9 on the BITSEA ASD + Dysregulation scale produced adequate sensitivity but poor specificity. Considering that reliable diagnoses of ASD often do not occur until children are 18 to 24 months (Kim & Lord, 2012; Lord et al., 2006), it is not surprising that year 1 scores are inconclusive. However, the BITSEA may be effective as a way of identifying children at 12 months that need close monitoring over the following year. For example, children without ASD demonstrate significant gains in both verbal and nonverbal communication from 14 to 24 months while children with ASD make gains only in the verbal domain (Charman et al., 2002). Although many children scoring “at risk” on the BITSEA in year 1 will not meet criteria for ASD, monitoring advancements in the use of nonverbal communication over the next year can help practitioners make decisions about whether to refer for further evaluation at that time. Moreover, it is possible that at risk for ASD in year 1 will indicate risk for other related developmental disorders. Future research should explore long-term diagnostic outcomes for children evaluated with the BITSEA to examine how early risk scores are related to later symptomology.

Along these lines, the current study utilized a combined ASD + Dysregulation score. It is important to note that the addition of the Dysregulation scale only slightly improved accuracy in detecting ASD risk; however, it may facilitate identification of children with dysregulation disorder or more severe ASD symptomology. Practitioners may wish to consider using the single ASD scale to simplify the screening process for ASD; however, knowledge of dysregulation symptoms (either from the BITSEA or through other evaluation processes) should not be ignored. The presence and severity of dysregulation symptoms may indicate presence of a dysregulation disorder, early signs of ASD, or other difficulties that require monitoring or follow-up evaluation.

Not surprisingly, the ASD scale yielded better accuracy in year 2 for both Gardner et al. (2013) and the current study. These findings suggest that the BITSEA ASD scale adequately identifies 2-year-olds at risk for ASD and it can be utilized in an African American population. An optimal cutpoint of 6 yielded good sensitivity (80.4) and specificity (81.2). Interestingly, these statistics were slightly better than the score of -12 identified by Gardner et al (2013). Sensitivity of .90 has also been suggested as appropriate for screeners (Jenkins, Hudson, & Johnson, 2007), as the value of identifying a child with ASD to facilitate receipt of needed services is greater than the cost of evaluating a child who ultimately does not need treatment. Hence, a cutpoint of 4 with nearly excellent sensitivity (89.3%) may be preferred. Practitioners utilizing the BITSEA are urged to use their professional judgment and consider additional factors (e.g., severity of symptoms, parent concerns, maternal stress, family history) as a single BITSEA cut point should not be used in isolation to indicate ASD risk.

One notable strength of this study was the participation of a large, African American sample with approximately two-thirds lower SES (i.e., based on health insurance status of Medicaid or no insurance). Research indicates that ASD affects racial and economic groups similarly; however, rates of diagnosis are lower for African American preschoolers and children from lower SES homes (Biao, 2012; Liptak et al., 2008). Therefore, it is important to examine the utility of ASD screeners in this specific population. Previous research investigating the BITSEA has typically utilized primarily White, higher SES samples. Results of the current study indicated that reliability statistics and relationships with the M-CHAT were lower than expected compared to previous research using more diverse samples (Gardner et al., 2013; Briggs-Gowan et al., 2004; Karabekiroglu et al., 2010). Considering the weaker relationships observed in the current sample and disparities in ASD diagnosis within the African American and lower SES populations, it is possible that these issues may be related. Although it is beyond the scope of this paper to identify why these lower values were found in our sample, it is possible that either the manifestation of ASD characteristics or maternal perception and reporting of these behaviors vary somewhat across racial and/or economic groups. Researchers and practitioners may need to consider these cultural differences when evaluating and interpreting ASD screening measures.

Maternal Factors

A number of maternal factors (e.g., age, education, parenting stress, and depression) have been associated with ASD diagnosis. Hence, the second aim of this study was to examine the usefulness of maternal variables in facilitating identification of

ASD risk status. It was hoped that the addition of maternal variables particularly related to parenting stress and psychological distress, could improve the accuracy of ASD screening in toddlers.

One of the most consistent maternal correlates of ASD is advancing maternal age (e.g., Bilder et al., 2009; Mandell et al., 2009; Shelton et al., 2010). More specifically, the risk increases significantly for mothers 35 years and older. The current study did not find a significant association between maternal age and ASD risk status. It is possible that the current sample did not have sufficient variability to detect such differences because the vast majority (70%) of mothers were between 21 and 31 years of age and only 9.1% were 35 years or older. Higher education level also has been associated with increased risk for ASD diagnosis (Biao, 2012; Liptak et al., 2008). Consistent with this research, the current study found that maternal education level was associated with ASD. Moreover, participants with private insurance were more than twice as likely as those without insurance or on Medicaid to be at risk for ASD. A potential explanation for these findings in previous research is that parents with higher educational attainment or economic means may be more aware of ASD and have access to better healthcare services. Therefore, they would be more likely to pursue evaluation and services for their children. It is important to note; however, that mothers were recruited for this study while pregnant and without concern for their children's ASD risk.
Considering these things, higher SES (as indicated by education or health insurance status) may be a risk factor for ASD risk.

Consistent with previous research of maternal stress (e.g., Davis & Carter, 2008), mothers of toddlers identified at risk for ASD at year 2 demonstrated significantly higher levels of parenting stress and psychological distress at both year 1 and year 2. Moreover, higher levels of parenting stress and psychological distress significantly increases the probability that children are at risk for ASD. This finding suggests that a relationship between children's ASD symptoms and maternal stress begins at a very early age. On the other hand, the correlations between these factors are weak and the increase in probability is relatively low when trying to utilize the scores to inform risk status. In year 1, the AUC increased from .67 for the BITSEA ASD + Dysregulation to .86 for the final model. Much of this change was explained by the health insurance variable, which indicated children of higher economic means were more than twice as likely to be at moderate to high risk for ASD. In year 2, the AUC only increased from .86 for the BITSEA ASD scale to .87 for the final model. Although important for theoretical purposes, specific scores on measures of parenting stress or psychological distress are unlikely to help guide pediatricians' decisions regarding ASD risk. However, recognizing a higher level of parenting stress or psychological distress, along with the presence of ASD symptomology, may inform practitioners' next steps.

Limitations and Future Directions

Several limitations of this study should be considered when interpreting and generalizing the findings. The goal of this study was to examine the utility of the BITSEA in an urban, African American sample with predominantly low income mother-child dyads. Due to this design, findings should be interpreted within this context and generalization to other populations limited. Notably, weaker psychometric properties were found with the current sample compared to more diverse samples. Further research is needed to purposefully examine cultural differences in the psychometric properties of the BITSEA and other social-emotional ratings.

Perhaps the most significant limitation of the study was the use of the M-CHAT as the outcome measure to indicate ASD risk. Although the M-CHAT may be widely used as a measure of ASD risk, it would be advantageous to examine the classification accuracy of the BITSEA with formal diagnosis of ASD as the outcome. Considering the findings from the present study, future research should consider longitudinal outcome data that includes diagnostic criteria for ASD. Along these lines, the latest revisions of ASD criteria (APA, 2013), may influence research examining ASD screeners. The M-CHAT was created and standardized using the previous diagnostic criteria for ASD, hence, its indicator of ASD risk is based on outdated information. Future research on the BITSEA should consider using the updated diagnostic criteria for ASD as the outcome variable.

Furthermore, maternal factors are clearly associated with ASD risk; however, it is still not clear how to utilize this information to facilitate identification of those children at risk. As symptoms of ASD stabilize around age 2, the current study found that the BITSEA ASD scale at 2 years provided adequate identification of children at risk; however, predictive validity of 1 year ratings was poor. Additional research is needed to improve identification of 1 year olds at risk for ASD. Moreover, the results regarding economic status are surprising given the sample. Higher maternal education and private insurance were associated with greater risk for ASD even though participants were drawn from a community population. Therefore, greater awareness of ASD symptoms and access to better healthcare were less likely to influence risk status. Future studies are needed to more closely examine the relationship of economic status, particularly maternal education and health insurance, as they contribute to ASD risk.

Summary

Overall, this study validated and extended the research base of ASD screening in toddlers. Results indicated that the BITSEA ASD scale can be used to accurately identify African American 2 year olds at risk for ASD. As expected, utilizing the BITSEA with 1-year-olds to identify children in need of further evaluation proved more difficult. However, the ASD and ASD + Dysregulation scales can be used to identify 1 year olds in need of close monitoring and follow-up screening at 2 years. Moreover, pediatricians can easily integrate the BITSEA into multiple well child exams throughout toddlerhood in order to monitor progress and ASD risk. Although parenting stress and psychological distress were significantly associated with ASD risk, incorporating these factors did not practically improve the classification accuracy of the BITSEA ASD scale at year 1. Utilizing parenting stress and psychological distress scores may not be feasible for pediatricians; however, the presence of these maternal characteristics may increase concerns regarding the need for follow-up evaluation. Finally, maternal education and income levels proved more important in identifying risk than expected. Further research examining the utility of the BITSEA is needed to explore the diagnostic accuracy of these scales with confirmed ASD diagnoses. Special attention to cultural differences should also be considered.

References

Abidin, R. R. (1995a). *Parenting Stress Index: Professional Manual* (3rd ed.). Odessa, FL: Psychological Assessment Resources, Inc.

Abidin, R. R. (1995b). *Parenting Stress Index, Third Edition*. Odessa, FL: Psychological Assessment Resources, Inc.

Achenbach, T. M. (2001). *Child Behavior Checklist.* Burlington, VT: University of Vermont, Department of Psychiatry.

Ainsworth, M. D., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychological Association, 46*, 333-341.

American Academy of Pediatrics. (2009). *Caring for your baby and young child: Birth to age 5* (5th ed.). S. P. Shelov & T. R. Altman (Eds.). Random House Publishing Group.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.

Baillargeon, R., Zoccolillo, M., Keenan, K., Cote, S., Perusse, D., Wu, H.,... & Tremblay, R. (2007). Gender differences in physical aggression: A prospective population-based survey of children before and after 2 years of age. *Developmental Psychology, 43*(1), 13-26. doi: 10.1037/0012-1649.43.1.13

Baio, J. (2012). Prevalence of Autism Spectrum Disorders: Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 61*(3). Centers for Disease Control and Prevention.

Baker, B. L., Blacher, J., Crnic, K. A., & Edelbrock, C. (2002). Behavior problems and parenting stress in families of three-year-old children with and without developmental delays. *American Journal on Mental Retardation, 107,* 433-444.

Baker, B. L., McIntyre, L. L., Blacher, K., Crnic, C., Edelbrock, C., & Low, C. (2003). Pre-school children with and without developmental delay: Behaviour problems and parenting stress over time. *Journal of Intellectual Disability Research, 47,* 217-230.

Baldessarini, R. J., Finkelstein, S., & Arana. G. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry, 40,* 569-573.

Baranek, G.T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. *Journal of Autism and Developmental Disorders, 29,* 213-224.

Bhatia, M. S., Dhar, N. K., Signhal, P. K., Nigam, V. R., Malik, S. C., & Mullick, D. N. (1990). Temper tantrums. Prevalence and etiology in a non-referral outpatient setting. *Clinical Pediatrics, 29,* 311-315.

Bilder, D., Pinborough-Zimmerman, J., Miller, J., & McMahon, W. (2009). Prenatal, perinatal, and neonatal factors associated with autism spectrum disorders. *Pediatrics, 123*, 1293-1300.

Blandon, A. Y., Calkins, S. D., & Keane, S. P. (2010). Predicting emotional and social competence during early childhood from toddler risk and maternal behavior. *Developmental Psychopathology, 2010, 22,* 119-132.

Briggs-Gowan, M. J. & Carter, A. S. (1998), Preliminary acceptability and psychometrics of the infant–toddler social and emotional assessment (ITSEA): A new adult-report questionnaire. *Infant Mental Health Journal, 19,* 422–445. doi:10.1002/(SICI)1097-0355(199824)19:4<422::AID-IMHJ5>3.0.CO;2-U.

Briggs-Gowan, M.J., & Carter, A.S. (2006). *BITSEA: Brief Infant-Toddler Social and Emotional Assessment Examiner's Manual.* San Antonio, TX: Harcourt Assessment.

Briggs-Gowan, M. J., & Carter, A. S. (2007). Applying the Infant-Toddler Social & Emotional Assessment (ITSEA) and Brief-ITSEA in early intervention. *Infant Mental Health Journal, 28*(6), 564-583.

Briggs-Gowan, M. J., Carter, A. S., Bosson-Heenan- J., Guyer, A. E., & Horwitz, S. M. (2006). Are infant-toddler social-emotional and behavioral problems transient? *Journal of American Academy of Child and Adolescent Psychiatry, 45*(7) 849-858.

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: Screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology, 29,* 495-515.

Briggs-Gowan, M. J., Carter, A. S., Skuban, E. M., & Horwitz, S. M. (2001). Prevalence of social-emotional and behavioral problems in a community sample of 1- and 2-year-old children. *Journal of American Academy of Child and Adolescent Psychiatry, 40*(7), 811-819.

Carter, A. S., & Briggs-Gowan, M. J. (2005). *Brief infant toddler social emotional assessment (BITSEA).* San Antonio, TX: Harcourt Assessment, Inc.

Carter, A. S., Garrity-Rokous, F. E., Chazan-Cohen, R. C., Little, C., & Briggs-Gowan, M. J. (2001). Maternal depression and comorbidity: Predicting early parenting attachment security, and toddler social-emotional problems and competencies. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(1), 18-26.

Centers for Disease Control and Prevention. (2014). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *Morbidity and Mortality Weekly Report Surveillance Summary*, 63, (SS-02), 1-21. http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm?s_cid=ss6302a1_w

Charman, T., & Baird, G. (2002). Practitioner Review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry, 43,* 289-305.

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L. E., & Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 62-72.

Cotton, S. M., & Richdale, A. L. (2010). Sleep patterns and behaviour in typically developing children and children with autism, Down syndrome, Prader-Willi syndrome, and intellectual disability. *Research in Autism Spectrum Disorders, 4*, 490-500.

Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J.,... Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry, 40,* 719-732.

Croen, L. A., Grether, J. K., & Selvin, S. (2002). The changing prevalence of autism in California. *Journal of Autism and Developmental Disorders, 32,* 217–224.

Croen, L. A., Najjar, D. V., Fireman, B., & Grether, J. K. (2007). Maternal and paternal age and risk of autism spectrum disorders. Archives of Pediatric and Adolescent Medicine, 161, 334-340.

Cummings, E. M., & Davies, P. T. (1994). Maternal depression and child development. *Journal of Child Psychology and Psychiatry, 35,* 73-112.

Davis, N. O., & Carter, A. S. (2008). Parenting stress in mothers and fathers of toddlers with autism spectrum disorders: Associations with child characteristics. *Journal of Autism and Developmental Disorders, 38,* 1278-1291.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., & Greenson, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The early start denver model. *Pediatrics, 125*, 17-23.

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child & Adolescent Psychiatry, 7,* 131-136.

Derogatis, L. R. (1993a). *Brief Symptom Inventory.* Baltimore, MD: Clinical Psychometric Research.

Derogatis, L. R. (1993b). *Brief Symptom Inventory (BSI). Administration, scoring and procedures manual.* Minneapolis: NCS Pearson, Inc.

Dodds, L., Fell, D. B., Shea, S., Armson, B. A., Allen, A. C., & Bryson, S. (2011). The role of prenatal, obstetric and neonatal factors in the development of autism. *Journal of Autism and Developmental Disorders*, *41*(7), 891-902.

Dominick, K. C., Davis, N. O., Lainhart, J., Tager-Flusberg, H., & Folstein, S. (2007). Atypical behaviors in children with autism and children with a history of language impairment. *Research in Developmental Disabilities, 28*(2), 145-162.

Dosreis, S., Weiner, C.L., Johnson, L., & Newschaffer, C.J. (2006). Autism spectrum disorder screening and management practices among general pediatric providers. *Journal of Developmental and Behavior Pediatrics, 27,* s88-s94.

Dumas, J., Wolf, L., Fisher, S., & Culligan, J. (1991). Parenting stress, child behavior problems, and dysphoria in parents of children with autism, Down syndrome, behavior disorders, and normal development. *Exceptionality, 2*, 97-110.

Dumont-Mathieu, T., & Fein, D. (2005). Screening for autism in young children: The Modified Checklist for Autism in Toddlers (M-CHAT) and other measures. *Mental Retardation & Developmental Disabilities Research Reviews, 11*(3), 253-262.

Estes, A., Munson, J., Dawson, G., Koehler, E., Zhou, X., & Abbott, R. (2009). Parenting stress and psychological functioning among mothers of preschool children with autism and developmental delay. *Autism, 13,* 375–387.

Gardner, L., Murphy, L., Campbell, J. M., Tylavsky, F., Palmer, F. B., & Graff, J. C. (2013). Screening accuracy for risk of autism spectrum disorder using the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). *Research in Autism Spectrum Disorders, 7*(5), 591-600.

Glascoe, F. P. (1991). Developmental screening: Rationale, methods, and application. *Infants & Young Children*, *4*(1), 1-10.

Gomez, C.R., & Baird, S. (2005). Identifying early indicators for autism in self-regulation difficulties. *Focus on Autism and Other Developmental Disabilities, 20*(2), 106-116.

Hartley, S. L., Sikora, D. M., & McCoy, R. (2008). Prevalence and risk factors of maladaptive behavior in young children with autistic disorder. *Journal of Intellectual Disabilities, 52*(10), 819-829.

Hastings, R. P. (2003). Child behavior problems and partner mental health as correlates of stress in mothers and fathers of children with autism. *Journal of Intellectual Disability Research, 47,* 231-237.

Hastings, R. P., & Johnson, E. (2001). Stress in UK families conducting intensive home-based behavioral intervention for their young child with autism.

Howlin, P., & Moore, A. (1997). Diagnosis in autism: a survey of over 1200 patients in the UK. *Autism, 1*, *135-162.*

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at risk readers in a response to intervention framework. *School Psychology Review, 36*(4), 582-600.

Johnson, C. P., & Meyers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120,* 1183-1215.

Karabekiroglu, K., Briggs-Gowan, M. J., Carter, A. A., Rodopman-Armond, A., & Akbasa, S. (2010). The clinical validity and reliability of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA). *Infant Behavior and Development, 33,* 503-509.

Kim, S. H., & Lord, C. (2012). New autism diagnostic interview-revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders, 42*(1), 82-93, DOI: 10.1007/s10803-011-1213-1

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L.,... Fein, D. (2008). The Modified Checklist for Autism in Toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(5), 827–839.

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Koegel, R., Koegel, L., & Surratt, A. (1992). Language intervention and disruptive behavior in preschool children with autism. *Journal of Autism and Developmental Disorders, 22,* 141-153.

Krug, D. A., Arick, J. R., & Almond, P. J. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, 21*, 221–229.

Kuban, K., O'Shea, T., Allred, E., Tager-Flusberg, H., Goldsten, D., & Leviton, A. (2009). Positive screening on the Modified Checklist for Autism in Toddlers (M-CHAT) in extrememly low gestational age newborns. *Pediatrics, 154*(4), 535-540.

Kuusikko-Gauffin, S., Pollock-Wurman, R., Mattila, M. L., Jussila, K., Ebeling, H., Pauls, D., & Moilanen, I. (2013). Social anxiety in parents of high-functioning children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *43*(3), 521-529.

Landa, R., Holman, K., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry, 64,* 853-864.

Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: A review of research in the last decade. *Psychological Bulletin, 137,* 562-593.

Limperopoulos, C., Bassan, H., Sullivan, N., Soul, J., Robertson, R., Moore, M... du Plessis, Adre. (2008). Positive screening for autism in ex-preterm infants: prevalence and risk factors. *Pediatrics, 121*(4), 758-765.

Liptak, G. S., Benzoni, L. B., Mruzek, D. W., Nolan, K. W., Thingvoll, M. A., Wade, C. M., & Fryer, G. E. (2008). Disparities in Diagnosis and Access to Health Services for Children with Autism: Data from the National Survey of Children's Health. *Journal of Developmental and Behavioral Pediatrics, 29*, 152-160.

Loh, A. Soman, T. Brian, J., Bryson, S., Roberts, W., Szatmari, P.,... Zwaigenbaum, L. (2007). Stereotyped motor behaviors associated with autism in high-risk infants: A pilot videotapes analysis of a sibling sample. *Journal of Autism and Developmental Disorders*, *37*, 25-36.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36,* 1365-1382.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism From 2 to 9 Years of Age. *Archives of General Psychiatry, 6*3(6), 694-701.

Mandell, D. S., Ittenbach, R. F., Levy, S. E., & Pinto-Martin, J. A. (2007). Disparities in diagnosis received prior to a diagnosis of autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(9), 1795-1802.

Mandell, D. S., Listerud, J., Levy, S. E., & Pinto-Martin, J. A. (2002). Race differences in the age at diagnosis among Medicaid-eligible children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(12), 1447–1453.

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics, 116*(6), 1480–1486.

Mandell, D. S., Wiggins, L. D., Carpenter, L. A., Daniels, J., DiGuiseppi, C., Durkin, M. S.,... Kirby, R. S. (2009). Racial/Ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health, 99*(3), 493-498.

Mathiesen, K. S., & Sanson, A. (2000). Dimensions of early childhood behavior problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology*, *28,* 15–31.

Matson, J. L., Mahan, S., Sipes, M., & Kozlowski, A. M. (2010). Effects of symptoms of comorbid psychopathology on challenging behaviors among atypically developing infants and toddlers assessed with the Baby and Infant Screen for Children with Autism Traits (BISCUIT). *Journal of Mental Health Research in Intellectual Disabilities, 3*(3), 164-176.

Matson, J. L., Wilkins, J., & Gonzalez, M. (2008). Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early. *Research in Autism Spectrum Disorders, 2,* 75-84.

Matson, J. L., Wilkins, J., & Macken, J. (2009). The relationship of challenging behaviors to severity and symptoms of autism spectrum disorders. *Journal of Mental Health Research in Intellectual Disabilities, 2*, 29-44.

Mesman, J., & Koot, K. M. (2001). Early preschool predictors of preadolescent internalizing and externalizing DSM-IV diagnoses. *Journal of American Academy of Child and Adolescent Psychiatry, 40,* 1029-1036.

Nachshen, J. S., Garcin, N., & Minnes, P. (2005). Problem behavior in children with intellectual disabilities: Parenting stress, empowerment, and school services. *Mental Health Aspects of Developmental Disabilities, 8,* 105-114.

Nadig, A. S., Ozonoff, S., Young, G. S., Rozga, A., Sigman, M., & Rogers, S. J. (2007). A prospective study of response to name in infants at risk for autism. *Archives of Pediatric and Adolescent Medicine, 161,* 378-383.

Nicholas, J. S., Charles, J. M., Carpenter, L. A., King, L. B., Jenner, W., & Spratt, E. G. (2008). Prevalence and characteristics of children with autism spectrum disorders. *Annals of Epidemiology, 18,* 130-136.

Nunnally, J. (1978). *Psychometric methods.* New York: McGraw.

Olsson, M. B., & Hwang, C. P. (2001). Depression in mothers and fathers of children with intellectual disability. *Journal of Intellectual Disability Research, 45*(6), 535-543.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: a study of first birthday home videotapes. *Journal of Autism and Developmental Disorders, 24,* 247-257.

Osterman, K., & Bjonkquist, K. (2010). A cross-sectional study of onset, cessation, frequency, and duration of children's temper tantrums in a nonclinical sample. *Psychological Reports, 106,* 448-454.

Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism, 12*(5), 457–472.

Pandey, J., Verbalis, A., Robins D., Boorstein, H., Klin, A. Babitz, T.,... Fein, D. (2008). Screening for autism in older and younger toddlers with the modified checklist for autism in toddlers. *Autism, 12*, 513-535.

Patzold, L., Richdale, B. J., & Tonge, B. J. (1998). An investigation into sleep characteristics of children with autism and Asperger's disorder. *Journal of Pediatric Child Health, 34,* 528-533.

Peduzzi, P., Concato, J., Feinstein, A., & Holford, T. (1995). Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology, 48*(12), 1503-1510.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturney, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders, 5*(1), 60-69.

Pierce, K., Carter, C., Weinfield, M., & Desmond, J. (2011). Detecting, studying, and treating autism early: the 1 year well baby check-up. *Pediatrics, 159*, 458-465.

Potegal, M., & Davidson, R. J. (2003). Temper tantrums in young children: 1. Behavioral composition. *Developmental and Behavioral Pediatrics, 24,* 140-147.

Richdale, A., & Prior, M. R. (1995). The sleep/wake rhythm in children with autism. *European Child and Adolescent Psychiatry, 4,* 175-186.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism, 12*(5), 537–556.

Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for Autism Spectrum Disorders: Update on the Modified Checklist for Autism in Toddlers and other measures. *Developmental and Behavioral Pediatrics, 27,* 111-119.

Robins D. L., Fein D., & Barton M. L. (1999). The Modified-Checklist for Autism in Toddlers. Self-published at https://www.m-chat.org

Robins, D., Fein, D., Barton, M., & Green, J. A. (2001). The Modified-Checklist for Autism in Toddlers (M-CHAT): an initiation investigation in the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31,* 131-144.

Rogers, S. J. (2009). What are infant siblings teaching us about autism in infancy?. *Autism Research, 2*, 125–137. doi:10.1002/aur.81

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology, 37*(1), 8-38.

Rojahn, J., Matson, J. L., Mahan, S., Fodstad, J. C., Knight, C., Sevin, J. A., & Sharp, B. (2009). Cutoffs, norms, and patterns of problem behaviors in children with an ASD on the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT-Part 3). *Research in Autism Spectrum Disorders, 3,* 989-998.

Rosen, W. D., Adamson, L. B., & Bakeman, R. (1992). An experimental investigation of infant social referencing: Mothers' messages and gender differences. *Developmental Psychology, 28*(6), 1172-1178.

Sanders, J. L., & Morgan, S. B. (1997). Family stress and adjustment as perceived by parents of children with autism or down syndrome: Implications for intervention. *Child & Family Behavior Therapy, 19*(4), 15-32.

Shelton, J. F., Tancredi, D. J., & Hertz-Picciotto, I. (2010). Independent and dependent contributions of advanced maternal and paternal ages to autism risk. *Autism Research, 3*(1), 30-39.

Smith, T. B., Oliver, M. N., & Innocenti , M. S. (2001). Parenting stress in families of children with disabilities. *American Journal of Orthopsychiatry, 71*(2), 257-261.

Swets, J. A. (1998). Measuring the accuracy of diagnostic systems. *Science, 240,* 1285-1293.

Stone, W., Lee, E., Ashford, L., Brissie, J., Hepburn, S., Coonrod, E., & Weiss, B. H. (1999). Can autism be diagnosed accurately in children under three years? *Journal of Child Psychology and Psychiatry, 40,* 219-226.

Striano, T., & Rochat, P. (2000). Emergence of selective social referencing in infancy. *Infancy, 1*(2), 253-264.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston: Pearson.

Titelbaum, P., Teitelbaum, O., Nye, J., Fryman, J., & Maurer, R. G. (1998). Movement analysis in infancy may be useful for early diagnosis of autism. *Process of National Academy of Sciences U.S.A., 95,* 13982-13987.

Tomanik, S., Harris, G., & Hawkins, J. (2004). The relationship between behaviors exhibited by children with autism and maternal stress. J*ournal of Intellectual & Developmental Disability, 29*(1), 16–26.

Valicenti-McDermott, M., Hottinger, K., Seijo, R., & Shulman, L. (2012). Age at diagnosis of Autism Spectrum Disorders. *Journal of Pediatrics, 161,* 554-556.

Watt, N., Wetherby, A. M., Barber, A., & Morgan, L. (2008). Repetitive and stereotyped behavior sin children with autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 38*(8), 1518-1533.

Werner, E., & Dawson, G. (2005). Validation of the phenomenon of autistic regression using home videotapes. *Archives of General Psychiatry, 62,* 889-895.

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34,* 473-493.

Woods, J. J., & Wetherby, A. M. (2003). Early identification of and intervention for infants and toddlers who are at risk for autism spectrum disorder. *Language, Speech & Hearing Services in Schools, 34,* 180-193.

Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver. L., … Yirmira, N. (2009). Clinical Assessment and Management of Toddlers With Suspected Autism Spectrum Disorder: Insights From Studies of High-Risk Infants. *Pediatrics, 123*(5), 1383 -1391.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestation of autism in the first year of life. *International Journal of Developmental Neuroscience, 23,* 143-152.